(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) Publication number: 0 685 559 A1

(12) **EUROPEAN PATENT APPLICATION published in accordance with Art. 158(3) EPC**

(21) Application number: 94927813.9

(51) Int. Cl.6: C12N 15/00

(22) Date of filing: 30.09.94

(86) International application number: PCT/JP94/01627

(87) International publication number: WO 95/09234 (06.04.95 95/15)

(30) Priority: 30.09.93 JP 267884/93

(43) Date of publication of application: 06.12.95 Bulletin 95/49

(84) Designated Contracting States: AT BE CH DE DK ES FR GB GR IE IT LI LU MC NL PT SE

(71) Applicant: Japan Tobacco Inc.
4-12-62 Higashishinagawa
Shinagawa-ku,
Tokyo 140 (JP)

(72) Inventor: UEKI, Jun, Japan Tobacco Inc., Plant Breeding &
Genetics Research Laboratory,
700, Higashibara,
Toyoda-cho, Iwata-gun,
Shizuoka 438 (JP)
Inventor: MORIOKA, Shinji, Japan Tobacco Inc, Plant Breeding
& Genetics Research Laboratory,
700, Higashibara
Toyoda-cho, Iwata-gun,
Shizuoka 438 (JP)

(74) Representative: Reinhard, Horst, Dr.
Reinhard-Skuhra-Weise & Partner
Patentanwälte
Postfach 44 01 51
D-80750 München (DE)

(54) PLANT-DERIVED PHOSPHOLIPASE D GENE.

(57) A cloned DNA that codes for a plant-derived phospholipase D, and a cloned DNA that controls the expression of a plant-derived phospholipase D gene.

EP 0 685 559 A1

TECHNICAL FIELD

The present invention relates to a phospholipase D gene originated from a plant.

BACKGROUND ART

Phospholipase D (hereinafter also referred to as "PLD") is one of phospholipid-decomposing enzymes, which catalyzes, for example, the reaction of decomposing lecithin to liberate phosphatidic acid and choline. This enzyme is known to occur in plants, animals and microorganisms. PLD is used as a reagent for measuring phospholipids in blood, as well as for hydrolysis of phospholipids and for production of derivatives by a base-exchange reaction utilizing the reversible reaction by PLD.

Purification and partial purification of PLDs originated from plants have been reported. That is, purification of rice PLD (Men Hui Lee, 1989, Phospholipase D of rice bran, I, purification and characterization, Plant Science 59, 25-33); partial purification of rice PLD (Katsumi TAKANO, Ikuzo KAMOI and Tetsujiro OBARA, 1987, About separation and purification, and properties of rice bran phospholipase D, J. Jpn. Soc. Food Sci. Technol., 34, 8-13 (1987)); purification of peanuts PLD (Michael Heller, Nava Mozes, Irena Peri and Eddie Maes, 1974, Phospholipase D from peanut seeds, IV, Final purification and some properties of the enzyme, Biochem. Biophys, Acta 369, 397-410); and purification of cabbage PLD (Romy Lambrecht and Renate Ulbrich-Hofmann, 1992, A facile purification procedure of phospholipase D from cabbage and its characterization, Biol. Chem. Hoppeseyler 373(2), 81-88) have been reported. However, amino acid sequences of plant PLDs have not been reported at all and genetical analyses thereof have also not been reported.

On the other hand, analyses of PLD genes of microorganisms and animals have been reported (Japanese Laid-open Patent Application (Kokai) No. 3-187382; Adrian L. M. Hodgson, Phillip Bird, and Ian T. Nisbet 1990, Cloning, nucleotide sequence, and expression in *Escherichia coli* of the phospholipase D gene from *Corynebacterium pseudotuberculosis*, Journal of Bacteriology 172, 1256-1261; and Japanese Laid-open Patent Application (Kokai) No. 5-76357).

If a PLD gene originated from a plant is available, the PLD originated from the plant may be produced in a large scale by a genetic engineering process, which is industrially advantageous. However, as mentioned above, since the amino acid sequence and DNA sequence of plant PLD have not been reported, it is hitherto impossible to genetically manipulate the PLD gene. Further, as mentioned above, although the amino acid sequences and DNA sequences of PLDs of microorganisms and animals have been reported, since homologies of sequences are not observed among the PLD genes of microorganisms or among the PLD genes of microorganisms and animals, it is difficult to isolate plant PLD gene based on the reported information. Further, since there is no information about plant PLD gene, an antisense DNA which suppresses expression of the plant PLD gene has not been obtained.

DISCLOSURE OF THE INVENTION

An object of the present invention is to provide a PLD gene originated from a plant. Another object of the present invention is to provide an antisense DNA which can suppress the expression of the above-mentioned PLD gene according to the present invention. Still another object of the present invention is to provide a DNA which regulates expression of the PLD gene originated from a plant.

After intensive study, the present inventors succeeded in isolation of rice PLD gene and in sequencing the PLD gene by purifying PLD from rice, determining the partial amino acid sequence thereof, screening a rice cDNA library using as a probe the PCR product obtained by using the oligonucleotides encoding the above-mentioned partial amino acid sequence, cloning the inserted gene of positive clones and sequencing the inserted gene. Further, the present inventors succeeded in obtaining a cDNA clone of maize PLD using as a probe the DNA encoding rice PLD, and in sequencing the maize PLD gene. The present inventors still further succeeded in isolating a genome DNA clone carrying the regulatory sequence of rice PLD gene and in sequencing it.

That is, the present invention provides a DNA encoding phospholipase D originated from a plant. The present invention also provides a DNA which regulates the expression of phospholipase D gene originated from a plant.

By the present invention, a PLD gene originated from a plant was cloned for the first time. By using the DNA according to the present invention, the PLD originated from a plant, which is industrially useful, can be produced in a large scale by a genetic engineering process. Further, by the present invention, a DNA which regulates expression of the PLD gene originated from a plant was cloned for the first time. By virtue of this,

expression of lipid-related genes may be suppressed, thereby modifying plant lipids.

BEST MODE FOR CARRYING OUT THE INVENTION

As mentioned above, the DNA according to the present invention encodes PLD originated from a plant. In the examples hereinbelow described, the DNA encoding rice PLD was isolated and sequenced. The deduced amino acid sequence encoded by the thus sequenced DNA is shown in SEQ ID NO. 2 in the Sequence Listing. The experimentally determined nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO. 2 is shown in SEQ ID NO. 1. Therefore, the amino acid sequence shown in SEQ ID NO. 1 is identical to the amino acid sequence shown in SEQ ID NO. 2. Further, in the examples hereinbelow described, a cDNA clone of maize PLD gene was isolated using as a probe the thus sequenced DNA encoding the rice PLD and the maize PLD gene was sequenced. The deduced amino acid sequence encoded by the thus sequenced DNA is shown in SEQ ID NO. 4 in the Sequence Listing. The experimentally determined nucleotide sequence encoding the amino acid sequence shown in SEQ ID NO. 4 is shown in SEQ ID NO. 3. Therefore, the amino acid sequence shown in SEQ ID NO. 3 is identical to the amino acid sequence shown in SEQ ID NO. 4.

Further, in the examples described below, rice cDNA library was screened using as a probe the thus obtained DNA encoding rice PLD and positive clones were isolated, followed by sequencing the inserted DNA to obtain a clone carrying an inserted DNA having the nucleotide sequence shown in SEQ ID NO. 5. In SEQ ID NO. 5, the 1876th base is "A" which is the first "A" in the translation initiation codon ATG of the PLD structural gene, and the region upstream of the "A" (i.e., 1-1875nt) is thought to be a regulatory region of the PLD gene. The amino acid sequence encoded by the region downstream of 1876nt is interrupted by intron.

The DNA according to the present invention may be obtained by, for example, the method described in the examples below in which rice or maize is used as the starting material. Alternatively, since the nucleotide sequence of the gene was disclosed by the present invention, the PLD gene may easily be prepared by PCR method using as primers a pair of oligonucleotides each of which hybridizes with the respective end of the DNA according to the present invention and using the genome DNA of rice or maize as a template.

A plant producing PLD may be obtained by inserting the DNA encoding PLD originated from a plant according to the present invention into a vector for plants by a conventional method and transforming a plant with the obtained recombinant vector by a conventional method.

By inserting the DNA according to the present invention into a vector for plants in the reverse direction, an antisense DNA of PLD gene may be obtained. By transforming a plant with such a recombinant vector, the mRNA which hybridizes with the mRNA formed by transcription of the PLD gene intrinsic to the plant is produced, so that expression of the PLD gene intrinsic to the plant may be suppressed. If this technique is applied to a rice plant, a rice plant in which expression of PLD gene is suppressed is obtained, so that the taste of the rice is improved.

The DNA regulating the PLD gene originated from a plant according to the present invention may be used in combination with the antisense DNA of the PLD gene. By this, the antisense DNA can be expressed at the same place as the place in which the PLD gene intrinsic to the plant is expressed.

It is well-known in the art that there are cases wherein the activity of an enzyme is retained even if the amino acid sequence of an enzyme is modified to a small extent, that is, even if one or more amino acids in the amino acid sequence are substituted or deleted, or even if one or more amino acids are added to the amino acid sequence. DNAs encoding the proteins having such modifications and having PLD activity are included within the scope of the present invention. That is, cloned DNAs encoding amino acid sequences having the same amino acid sequence as SEQ ID NO. 2 or SEQ ID NO. 4 except that one or more amino acids are substituted, deleted or added, which give the enzyme activity of PLD, are also included in the scope of the present invention. Similarly, DNAs having the same nucleotide sequence as SEQ ID NO. 5 except that one or more nucleotides are substituted, deleted or added, which regulate expression of the DNA encoding the amino acid sequence giving the enzyme activity of PLD are also included within the scope of the present invention.

Modification of DNA which brings about addition, deletion or substitution of the amino acid sequence encoded thereby can be attained by the site-specific mutagenesis which is well-known in the art (e.g., Nucleic Acid Research, Vol. 10, No. 20, p6487-6500, 1982). In the present specification, "one or more amino acids" means the number of amino acids which can be added, deleted or substituted by the site-specific mutagenesis.

Site-specific mutagenesis may be carried out by, for example, using a synthetic oligonucleotide primer complementary to a single-stranded phage DNA except that the desired mutation as follows. That is, using the above-mentioned synthetic oligonucleotide as a primer, a complementary chain is produced by a phage, and host bacterial cells are transformed with the obtained double-stranded DNA. The culture of the transformed bacterial cells is plated on agar and plaques are formed from a single cell containing the phage. Theoretically, 50% of the new colonies contain the phage having a single-stranded chain carrying the mutation and remaining 50% of the colonies contain the phage having the original sequence. The obtained plaques are then subjected to hybridization with a kinase-treated synthetic probe at a temperature at which the probe is hybridized with the DNA having exactly the same sequence as the DNA having the desired mutation but not with the original DNA sequence that is not completely complementary with the probe. Then the plaques in which the hybridization was observed are picked up, cultured and the DNA is collected.

In addition to the above-mentioned site-specific mutagenesis, the methods for substituting, deleting or adding one or more amino acids without losing the enzyme activity include a method in which the gene is treated with a mutagen and a method in which the gene is selectively cleaved, a selected nucleotide is removed, added or substituted and then the gene is ligated.

Examples

The present invention will now be described more concretely by way of examples. However, the present invention is not restricted to the examples described below.

1. Purification of rice bran PLD

For the purification, a reference relating to the purification of rice bran PLD (Takano et al., J. Jpn. Soc. Food Sci. Technol., 34, 8-13 (1987)) was referred. The enzyme activity was measured by using phosphatidylcholine as the substrate and by quantifying the choline generated by the enzyme reaction by the choline oxidase method (Imamura et al., J. Biochem. 83, 677-680 (1978)). The enzyme reaction, however, was stopped by heat treatment at 95 °C for 5 minutes.

That is, one liter of hexane was added to 100 g of rice bran of *Oryza sativa* variety "Koshihikari" and the mixture was stirred overnight to defat the rice bran. Then 10 g of POLYCRAL AT (trademark: polyvinylpyrrolidone commercially available from GAF Chemical) and 500 ml of 10 mM Tris-HCl buffer (pH 7) containing 1 mM $CaCl_2$ and 5 mM 2-mercaptoethanol were added and the resulting mixture was stirred for one hour to extract enzyme. The extract was filtered through 8-ply gauze and the filtrate was centrifuged at 15,000 x g for 20 minutes. The intermediate layer was collected as a crude extract. The crude extract was treated with ammonium sulfate (65% saturation) and formed precipitates were collected by centrifugation (15,000 x g, 20 minutes). The precipitates were dissolved and dialyzed against the above-described buffer. After the dialysis, precipitates were removed by centrifugation to obtain a ammonium sulfate fraction.

The ammonium sulfate fraction was applied to a DEAE-Cellulose (commercially available from Whattman) column (2.0 x 10 cm) equilibrated with buffer A (10 mM Tris-HCl pH 7, 1 mM $CaCl_2$, 1 mM 2-mercaptoethanol). After washing the column with about 100 ml of buffer A containing 50 mM NaCl, elution was performed with 120 ml of buffer A having linear gradient of NaCl from 50 to 350 mM. PLD was eluted in the vicinity of 0.2 M of NaCl. The fractions exhibiting PLD activity were collected as an eluted solution (DEAE-Cellulose).

To the eluted solution (DEAE-Cellulose), 3 M of ammonium sulfate was added to a concentration of 1 M and the resultant was applied to a Phenyl Sepharose column (commercially available from PHARMACIA, 2.6 x 10 cm). Then the column was subjected to elution with 240 ml of buffer A having a gradient of ammonium sulfate from 1.0 to 0 M. PLD was eluted in the vicinity of 0.1 M of ammonium sulfate. Fractions exhibiting PLD activity were collected and dialyzed against buffer A to obtain an eluted solution (Phenyl Sepharose).

The eluted solution (Phenyl Sepharose) was applied to a Mono Q column (anion-exchange column, commercially available from PHARMACIA, 16 x 10 cm) and then elution was carried out with 150 ml of buffer A having a gradient of NaCl from 50 - 350 mM. PLD was eluted at NaCl concentrations from 210 mM to 235 mM. The fractions exhibiting PLD activity were collected and dialyzed against buffer A to obtain an eluted solution (Mono Q 1st).

The eluted solution (Mono Q 1st) was concentrated to 0.5 ml by centrifugal ultrafiltration and the resultant was applied to Superose 6 column (commercially available from PHARMACIA, 1.0 x 30 cm) equilibrated with buffer A containing 0.1 M NaCl, and the column was subjected to elution with the same buffer. The molecular weight of the PLD was estimated as 78 kDa. The fractions exhibiting PLD activity

were collected as an eluted solution (Superose 6).

To the eluted solution (Superose 6), 2.5 ml of 40% CARRIER AMPHOLITE (commercially available from PHARMACIA, pH 4.0-6.0) and distilled water were added to a final volume of 50 ml. The resultant was then subjected to isoelectric focusing using Rotofore (commercially available from BIORAD). The electrophoresis was carried out at 2°C under a constant power of 12W for 4 hours. PLD activity was observed in the vicinity of pH4.9. The fractions exhibiting PLD activity were collected and the obtained solution was dialyzed against buffer A to obtain an isoelectric focusing fraction.

The isoelectric focusing fraction was applied to a Mono Q column (commercially available from PHARMACIA, 0.5 x 5 cm) and elution was carried out using buffer A having a linear gradient of NaCl from 50 mM to 350 mM. PLD was eluted in the vicinity of 210 mM and 235 mM NaCl. The two fractions exhibiting PLD activity were collected as eluted solutions (Mono Q 2nd-I and Mono Q 2nd-II).

Purities of the eluted solutions (Mono Q 2nd-I and Mono Q 2nd-II) were determined by SDS-polyacrylamide gel electrophoresis (Laemmli (1970)). After the electrophoresis, the gels were stained with COOMASSIE BRILLIANT BLUE R250. With either eluted solution, a main band was observed at the position of molecular weight of 82 kDa. A single band was observed with eluted solution (Mono Q 2nd-II).

By the above-described purification, eluted solutions (Mono Q 2nd-I and Mono Q 2nd-II) were 380-fold and 760-fold purified, respectively, based on the crude extract.

The two fractions were analyzed for properties of the enzymes contained therein. The results are shown in Table 1. The buffer solutions used for measuring the optimum pH were sodium acetate (pH4-6), MES-NaOH (pH5.5-7.0) and Tris-HCl (pH7-9), all of which had a concentration of 100 mM. The pH stability means the range in which decrease in activity was not observed after leaving the enzyme to stand at 25°C for 30 minutes. The temperature stability was measured by measuring remaining activity after leaving the enzyme to stand at 4°C, 25°C, 37°C or 50°C for 30 minutes. The substrate specificity was measured at a substrate concentration of 5 mM and is expressed as a relative activity taking the enzyme activity to phosphatidylcholine as 100.

Table 1

	Mono Q 2nd-I	Mono Q 2nd-II
Km Value	0.29 mM	0.29 mM
Optimum pH	6	6
pH Stability	7-8	7-8
Temperature Stability	4-37°C	4-37°C
Ca^{2+} Dependency	20 mM or more	20 mM or more
Substrate Specificity		
Phosphatidylcholine	100	100
Lysophosphatidylcholine	13	12
Sphingomyelin	6	4

2. Proof that Purified Protein is PLD

In the same manner as in the determination of the purity, eluted solutions (Mono Q 2nd-I and Mono Q 2nd-II) were separately subjected to SDS-polyacrylamide gel electrophoresis and each of the gels was transcribed to a PVDF membrane (commercially available from MILLIPORE), followed by staining the membrane. The band corresponding to the protein having a molecular weight of 82 kDa was cut out and amino acid sequence of the N-terminal region of the protein was determined. Amino acid sequence up to 10th amino acid residue was able to be determined for both proteins, and both proteins had the same sequence as follows:

Val Gly Lys Gly Ala Thr Lys Val Tyr Ser

Although the relationship between the proteins of 82 kDa which existed in two fractions having the enzyme activity was unknown, at least the homology of the amino acid sequences thereof was thought to be high. Thus, it was thought that there was no problem even if a mixture of these fractions was used as an antigen for producing an antibody.

A mixture of the eluted solutions (Mono Q 2nd-I and Mono Q 2nd-II) was subjected to SDS-polyacrylamide gel electrophoresis employing 7.5% acrylamide and the gel was stained with COOMASSIE

BRILLIANT BLUE R250. The band containing the protein of 82 kDa was cut out and the proteins were recovered by electroelution (25 mM Tris, 192 mM glycine, 0.025% SDS, 100V, 10 hours). After removing SDS by electrodialysis (15 mM ammonium bicarbonate, 200V, 5 hours), the resultant was freeze-dried. The electroelution and electrophoresis were performed by using BIOTRAP (commercially available from SCHLEICHER & SCHUELL).

Rabbits were immunized with the protein of 82 kDa highly purified by the above-described method. Immunization was performed by administering 50 µg of the protein per time at 7 days' interval. Using blood serum before the immunization and after the third immunization, immunotitration was performed. That is, PLD solution containing 8.6 x 10^{-3} units of the enzyme, 0 to 50 µl of the serum before the immunization or after the third immunization, 50 µl of 250 mM Tris-HCl (pH 7.0), 5 µl of 50 mM $CaCl_2$, 50 µl of 0.2% TRITON X-100 (trademark) and balance of water were mixed to a final volume of 250 µl and the mixture was left to stand at room temperature for 2.5 hours. To the resulting mixture, 200 µl of Protein A SEPHAROSE (commercially available from PHARMACIA) was added and the resultant was gently shaken at room temperature for 2 hours. The resultant was centrifuged (500 x g, 5 minutes) and enzyme activity of the supernatant was measured. Taking the enzyme activity when no serum was added as 100%, the enzyme activities measured when 20 µl or 50 µl of the serum before the immunization were 95% and 88%, respectively, while the enzyme activities measured when 20 µl or 50 µl of the serum after the third immunization were 75% and 30%, respectively. These results proved that the protein of 82 kDa is PLD.

3. Determination of Internal Amino Acid Sequence

The PLD protein was fragmented by fragmenting the protein in a gel (Cleveland et al., J. Biol. Chem., 252, 1102(1977)). The gel containing the PLD protein, which was cut out by the same method as described in 2 was inserted in a stacking gel well on a 15% acrylamide gel prepared for separation of peptides. *Staphylococcus aureus* V8 protease (commercially available from WAKO PURE CHEMICAL INDUSTRIES, LTD) in an amount of 1/10 of the PLD protein was overlaid and electrophoresis was started. The electrophoresis was interrupted when bromophenol blue reached the center of the stacking gel, and 30 minutes after, the electrophoresis was restarted. After the electrophoresis, the gel was transcribed to a PVDF membrane and the PVDF membrane was stained. Clear bands were observed at positions corresponding to molecular weights of 20, 14, 13, 11 and 10 kDa. The bands corresponding to molecular weights of 20, 14 and 13 kDa were cut out and amino acid sequences of the peptide fragments contained in the bands were determined by a protein sequencer. The sequences are as follows:

```
20 kDa:  Asn Tyr Phe His Gly Ser Asp Val Asn  ?  Val Leu
    ?  Pro Arg Asn Pro Asp Asp(Asp)  ?   ?  Ile
14 kDa:  Thr  ?  Asn Val Gln Leu Phe Arg Ser Ile Asp Gly
    Gly Ala Ala Phe Gly Phe Pro Asp Thr Pro Glu Glu Ala Ala
    Lys  ?  Gly Leu Val Ser Gly

13 kDa:  Ile Ala Met Gly Gly Tyr Gln Phe Tyr His Leu Ala
    Thr Arg Gln Pro Ala Arg Gly Gln Ile His Gly Phe Arg Met
    Ala Leu  ?  Tyr Glu His Leu Gly Met Leu  ?  Asp Val Phe
```

(In the sequences, "?" means the residue which could not be identified and the amino acid residue in parenthesis is one which may be another amino acid residue with a considerable probability.)

4. Preparation of cDNA Library of Rice Immature Seeds

Total RNAs were prepared from immature seeds 5 days after blossom by extracting RNAs by the SDS-phenol method and by precipitating the extract with lithium chloride. Poly(A)$^+$RNAs were prepared using OLIGOTEX-dT30 (commercially available from TAKARA SHUZO) in accordance with the instructions by the manufacturer. For cDNA cloning, cDNA SYNTHESIS SYSTEM PLUS (commercially available from AMERSHAM) and cDNA CLONING SYSTEM λgt10 (commercially available from AMERSHAM) were used. As the cloning vector, λZAPII vector (commercially available from STRATAGENE) was used and as the host cells, XL1-Blue was used.

5. Preparation of Probes

Oligonucleotides corresponding to the amino acid sequence of PLD were synthesized by a DNA synthesizer (commercially available from APPLIED BIOSYSTEMS). The sequences thereof and the amino acid sequences corresponding thereto are described below.

20KF: 5' AAYTAYTTYCAYGG 3'

20KR1: 5' RTCRTCRTCNGGRTT 3'

(wherein R represents purine bases, that is, A or G; Y represents pyrimidine bases, that is, T or C; and N represents G, A, T or C.)

20KF is a mixture of 32 types of oligonucleotides each of which encodes the amino acid sequence of

Asn Tyr Phe His Gly

that was found in the peptide having a molecular weight of 20 kDa, and 20KR1 is a mixture of 128 types of oligonucleotides each of which encodes the amino acid sequence of

Asn Pro Asp Asp(Asp)

that was found in the same peptide.

The cDNA synthesis reaction was carried out in a mixture of 10 ng of Poly(A)$^+$RNA, 0.3 μg of random hexamer (dN6), 10 U of RNase inhibitor (RNAGuard, commercially available from PHARMACIA), 1 mM each of dATP, dCTP, dGTP and dTTP, 1 x PCR buffer (commercially available from TAKAKA SHUZO), 50 mM of magnesium chloride and 100 U of reverse transcriptase (M-MuLV RTase, commercially available from BRL), the total volume of the reaction mixture being 10 μl. The reaction was carried out at 37°C for 30 minutes and the mixture was then heated at 95°C for 5 minutes, followed by retaining the resulting mixture in ice.

Polymerase chain reaction (PCR) was performed using the above-described cDNA as a template, and 20KF and 20KR1 as primers. The reaction was carried out using 10 μl of the cDNA synthesis reaction mixture, a mixture of 50 pmol each of the primers, 200 μM each of dATP, dCTP, dGTP and dTTP, 1 x PCR buffer (commercially available from TAKARA SHUZO) and 2.5 U of AmpliTaq DNA polymerase (commercially available from TAKARA SHUZO), the total volume of the reaction mixture being 50 μl. A cycle of 94°C for 1 minute/40°C for 1 minute/72°C for 2.5 minutes was repeated 30 times in DNA THERMOCYCLER (commercially available from PERKIN ELMER CETUS).

PCR product was separated on 2% agarose gel. Several fragments were detected by staining the gel with ethidium bromide. One of them had a size of 94 bp which is the expected size.

The PCR fragment was cut out from the gel and subcloned into pUC19 plasmid. The subcloned PCR fragment was sequenced by the dideoxy method using T7 sequencing kit (commercially available from PHARMACIA). Between the two primers, a nucleotide sequence encoding the expected amino acid sequence was observed. The nucleotide sequence between the two primers and the amino acid sequence encoded thereby are as follows:

```
C TCT GAC GTG AAC TGT GTT CTA TGC CCT CGC
 Ser Asp Val Asn Cys Val Leu Cys Pro Arg
```

Isotope ^{32}P (commercially available from AMERSHAM) was incorporated into the above-described oligonucleotide using DNA 5'-end labelling kit MEGALABEL (commercially available from TAKAKA SHUZO) to obtain a radioactive oligonucleotide probe.

6. Screening of PLD Gene-containing Clone

Using the above-described radioactive oligonucleotide as a probe, a cDNA library was screened. The hybridization solution was 0.5 M sodium phosphate buffer (pH 7.2) containing 7% SDS, 1 mM EDTA and 100 µg/ml of salmon sperm DNA, and hybridization was carried out at 45°C for 16 hours after adding the probe to this solution. The washing solution contained 0.3 M NaCl and 30 mM sodium citrate and washing was performed twice at 45°C for 20 minutes. Positive plaques were isolated and subcloned *in vivo* into pBluescript plasmid (commercially available from STRATAGENE) in accordance with the instruction of the manufacturer of λZAPII cloning vector. The nucleotide sequence was determined and the region encoding the internal amino acid sequence determined in 3 existed.

7. Sequencing of 5'-end Region

Since a clone containing the full-length cDNA could not be isolated, a DNA fragment containing the 5'-end region was prepared by RACE method (Edwards et al., Nucleic Acids Res., 19, 5227-5232(1991)). 5'-AmpliFINDER RACE Kit (commercially available from CLONETECH) was used in accordance with the instructions attached to the product. OligoDNAs were prepared and PCR was performed using the mRNA prepared by the method described in 4 as a template. The PCR product was subcloned into PCRII vector (commercially available from INVITROGEN) and sequenced by the dideoxy method. As a result, it was estimated that translation is initiated from the 182th nucleotide shown in SEQ ID NO. 1 because a termination codon exists upstream thereof by 36 bp.

8. Preparation of cDNA Clone Encoding PLD Originated from Maize

Maize cDNA clone was obtained by the method described below using the DNA encoding rice PLD as a probe.

Using suspended cultured cells established by culturing a callus derived from immature embryo of maize inbred Mo 17 (commercially available from MIKE BRAYTON SEEDS, INC.) in a liquid culture medium, a cDNA library was prepared by the method described in 4. However, λgt10 vector (commercially available from AMERSHAM) was used as the cloning vector and NM-514 (commercially available from AMERSHAM) was used as the host cells. Using the cDNA of rice PLD as a probe, positive plaques were isolated by the method described in 6. Phage DNA was prepared in accordance with the instructions by the manufacturer of the cloning vector. The phage DNA was digested with a restriction enzyme *Kpn* I and the resultant was subcloned into pBluescript plasmid. The nucleotide sequence was determined by the dideoxy method as described in 6.

9. Isolation of PLD Genome Clone Corresponding to PLD cDNA and Identification of Promoter Region

To isolate a genome DNA clone carrying the regulatory sequence of the PLD gene corresponding to the PLD cDNA sequenced in 6, which was cloned into pBluescript plasmid, a genome library of rice Koshihikari was prepared. This was carried out by partially digesting DNAs from leaves of Koshihikari with *Mbo* I, purifying fractions having sizes of 16-20 kb by sucrose gradient centrifugation, and by using lambda DASH II (commercially available from STRATAGENE) and GigapackII Gold (commercially available from STRATAGENE). Using the PLD cDNA clone as a probe, the genome library was screened. The screening was performed as described in 6. However, the hybridization was performed at 65°C for 16 hours, the washing solution was 0.5 x SSC containing 0.1% SDS and washing was performed twice at 65°C for 20 minutes. The nucleotide sequence of the hybridized genome clone was determined by the dideoxy method. As a result, a region homologous to the cDNA sequence determined in 6 existed.

The transcription initiation site was determined by the method described in 7. In the vicinity of the transcription initiation site, "TATA" consensus sequence box was observed. The ATG translation initiation site was determined as the upstream most ATG codon in the translation reading frame of the clone and as the ATG codon which is first accessible in the mRNA synthesized in rice.

A part of the DNA sequence of the genome clone hybridized with the cDNA clone is shown in SEQ ID NO. 5. In the genome DNA sequence, a reading frame starting from the ATG translation initiation codon, which overlaps with the corresponding cDNA sequence was identified. The promoter region is located upstream of the ATG translation initiation codon and starts immediately upstream thereof.

SEQ ID NO.:1

SEQUENCE LENGTH: 3040

SEQUENCE TYPE: NUCLEIC ACID

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE

ORGANISM: Oryza sativa

SEQUENCE DESCRIPTION

```
AGTCTCTCTT CTCCCGCAAT TTTATAATCT CGATCGATCC AATCTGCTCC CCTTCTTCTT   60
CTACTCTCCC CATCTCGGCT CTCGCCATCG CCATCCTCCT CTCCCTTCCC GGAGAAGACG  120
CCTCCCTCCG CCGATCACCA CCCGGTAGGG CGAGGAGGGA GCCAAATCCA AATCAGCAGC  180
C ATG GCG CAG ATG CTG CTC CAT GGG ACG CTG CAC GCC ACC ATC TTC GAG  229
  Met Ala Gln Met Leu Leu His Gly Thr Leu His Ala Thr Ile Phe Glu
    1               5                  10                  15
GCG GCG TCG CTC TCC AAC CCG CAC CGC GCC AGC GGA AGC GCC CCC AAG    277
Ala Ala Ser Leu Ser Asn Pro His Arg Ala Ser Gly Ser Ala Pro Lys
                20                  25                  30
TTC ATC CGC AAG TTT GTG GAG GGG ATT GAG GAC ACT GTG GGT GTC GGC    325
Phe Ile Arg Lys Phe Val Glu Gly Ile Glu Asp Thr Val Gly Val Gly
            35                  40                  45
AAA GGC GCC ACC AAG GTG TAT TCT ACC ATT GAT CTG GAG AAA GCT CGT    373
Lys Gly Ala Thr Lys Val Tyr Ser Thr Ile Asp Leu Glu Lys Ala Arg
        50                  55                  60
GTA GGG CGA ACT AGG ATG ATA ACC AAT GAG CCC ATC AAC CCT CGC TGG    421
Val Gly Arg Thr Arg Met Ile Thr Asn Glu Pro Ile Asn Pro Arg Trp
    65                  70                  75                  80
TAT GAG TCG TTC CAC ATC TAT TGC GCT CAT ATG GCT TCC AAT GTG ATC    469
Tyr Glu Ser Phe His Ile Tyr Cys Ala His Met Ala Ser Asn Val Ile
```

85 90 95

TTC ACT GTC AAG ATT GAT AAC CCT ATT GGG GCA ACG AAT ATT GGG AGG 517

Phe Thr Val Lys Ile Asp Asn Pro Ile Gly Ala Thr Asn Ile Gly Arg

100 105 110

GCT TAC CTG CCT GTC CAA GAG CTT CTC AAT GGA GAG GAG ATT GAC AGA 565

Ala Tyr Leu Pro Val Gln Glu Leu Leu Asn Gly Glu Glu Ile Asp Arg

115 120 125

TGG CTC GAT ATC TGT GAT AAT AAC CGC GAG TCT GTT GGT GAG AGC AAG 613

Trp Leu Asp Ile Cys Asp Asn Asn Arg Glu Ser Val Gly Glu Ser Lys

130 135 140

ATC CAT GTG AAG CTT CAG TAC TTC GAT GTT TCC AAG GAT CGC AAT TGG 661

Ile His Val Lys Leu Gln Tyr Phe Asp Val Ser Lys Asp Arg Asn Trp

145 150 155 160

GCG AGG GGT GTC CGC AGT ACC AAG TAT CCA GGT GTT CCT TAC ACC TTC 709

Ala Arg Gly Val Arg Ser Thr Lys Tyr Pro Gly Val Pro Tyr Thr Phe

165 170 175

TTC TCT CAG AGG CAA GGG TGC AAA GTT ACC TTG TAC CAA GAT GCT CAT 757

Phe Ser Gln Arg Gln Gly Cys Lys Val Thr Leu Tyr Gln Asp Ala His

180 185 190

GTC CCA GAC AAC TTC ATT CCA AAG ATT CCG CTT GCC GAT GGC AAG AAT 805

Val Pro Asp Asn Phe Ile Pro Lys Ile Pro Leu Ala Asp Gly Lys Asn

195 200 205

TAT GAA CCC CAC AGA TGC TGG GAG GAT ATC TTT GAT GCT ATA AGC AAT 853

Tyr Glu Pro His Arg Cys Trp Glu Asp Ile Phe Asp Ala Ile Ser Asn

210 215 220

GCT CAA CAT TTG ATT TAC ATC ACT GGC TGG TCT GTA TAC ACT GAG ATC 901

Ala Gln His Leu Ile Tyr Ile Thr Gly Trp Ser Val Tyr Thr Glu Ile

```
225                 230                 235                 240
ACC TTG GTT AGG GAC TCC AAT CGT CCA AAA CCT GGA GGG GAT GTC ACC    949
Thr Leu Val Arg Asp Ser Asn Arg Pro Lys Pro Gly Gly Asp Val Thr
                245                 250                 255
CTT GGG GAG TTG CTC AAG AAG AAG GCC AGT GAA GGT GTT CGG GTC CTC    997
Leu Gly Glu Leu Leu Lys Lys Lys Ala Ser Glu Gly Val Arg Val Leu
            260                 265                 270
ATG CTT GTG TGG GAT GAC AGG ACT TCA GTT GGT TTG CTA AAG AGG GAT   1045
Met Leu Val Trp Asp Asp Arg Thr Ser Val Gly Leu Leu Lys Arg Asp
        275                 280                 285
GGC TTG ATG GCA ACA CAT GAT GAG GAA ACT GAA AAT TAC TTC CAT GGC   1093
Gly Leu Met Ala Thr His Asp Glu Glu Thr Glu Asn Tyr Phe His Gly
    290                 295                 300
TCT GAC GTG AAC TGT GTT CTA TGC CCT CGC AAC CCT GAT GAC TCA GGC   1141
Ser Asp Val Asn Cys Val Leu Cys Pro Arg Asn Pro Asp Asp Ser Gly
305                 310                 315                 320
AGC ATT GTT CAG GAT CTG TCG ATC TCA ACT ATG TTT ACA CAC CAT CAG   1189
Ser Ile Val Gln Asp Leu Ser Ile Ser Thr Met Phe Thr His His Gln
                325                 330                 335
AAG ATA GTA GTT GTT GAC CAT GAG TTG CCA AAC CAG GGC TCC CAA CAA   1237
Lys Ile Val Val Val Asp His Glu Leu Pro Asn Gln Gly Ser Gln Gln
            340                 345                 350
AGG AGG ATA GTC AGT TTC GTT GGT GGC CTT GAT CTC TGT GAT GGA AGG   1285
Arg Arg Ile Val Ser Phe Val Gly Gly Leu Asp Leu Cys Asp Gly Arg
        355                 360                 365
TAT GAC ACT CAG TAC CAT TCT TTG TTT AGG ACA CTC GAC AGT ACC CAT   1333
Tyr Asp Thr Gln Tyr His Ser Leu Phe Arg Thr Leu Asp Ser Thr His
```

```
370                     375                     380
CAT GAT GAC TTC CAC CAG CCA AAC TTT GCC ACT GCA TCA ATC AAA AAG   1381
His Asp Asp Phe His Gln Pro Asn Phe Ala Thr Ala Ser Ile Lys Lys
385                 390                     395                 400
GGT GGA CCT AGA GAG CCA TGG CAT GAT ATT CAC TCA CGG CTG GAA GGG   1429
Gly Gly Pro Arg Glu Pro Trp His Asp Ile His Ser Arg Leu Glu Gly
                405                     410                     415
CCA ATC GCA TGG GAT GTT CTT TAC AAT TTC GAG CAG AGA TGG AGA AAG   1477
Pro Ile Ala Trp Asp Val Leu Tyr Asn Phe Glu Gln Arg Trp Arg Lys
            420                     425                     430
CAG GGT GGT AAG GAT CTC CTT CTG CAG CTC AGG GAT CTG TCT GAC ACT   1525
Gln Gly Gly Lys Asp Leu Leu Leu Gln Leu Arg Asp Leu Ser Asp Thr
        435                     440                     445
ATT ATT CCA CCT TCT CCT GTT ATG TTT CCA GAG GAC AGA GAA ACA TGG   1573
Ile Ile Pro Pro Ser Pro Val Met Phe Pro Glu Asp Arg Glu Thr Trp
    450                     455                     460
AAT GTT CAG CTA TTT AGA TCC ATT GAT GGT GGT GCT GCT TTT GGG TTC   1621
Asn Val Gln Leu Phe Arg Ser Ile Asp Gly Gly Ala Ala Phe Gly Phe
465                 470                     475                 480
CCT GAT ACC CCT GAG GAG GCT GCA AAA GCT GGG CTT GTA AGC GGA AAG   1669
Pro Asp Thr Pro Glu Glu Ala Ala Lys Ala Gly Leu Val Ser Gly Lys
                485                     490                     495
GAT CAA ATC ATT GAC AGG AGC ATC CAG GAT GCA TAC ATA CAT GCC ATC   1717
Asp Gln Ile Ile Asp Arg Ser Ile Gln Asp Ala Tyr Ile His Ala Ile
            500                     505                     510
CGG AGG GCA AAG AAC TTC ATC TAT ATA GAG AAC CAA TAC TTC CTT GGA   1765
Arg Arg Ala Lys Asn Phe Ile Tyr Ile Glu Asn Gln Tyr Phe Leu Gly
```

```
        515                      520                      525
AGT TCC TAT GCC TGG AAA CCC GAG GGC ATC AAG CCT GAA GAC ATT GGT   1813
Ser Ser Tyr Ala Trp Lys Pro Glu Gly Ile Lys Pro Glu Asp Ile Gly
    530                      535                      540
GCC CTG CAT TTG ATT CCT AAG GAG CTT GCA CTG AAA GTT GTC AGT AAG   1861
Ala Leu His Leu Ile Pro Lys Glu Leu Ala Leu Lys Val Val Ser Lys
545                     550                      555                 560
ATT GAA GCC GGG GAA CGG TTC ACT GTT TAT GTT GTG GTG CCA ATG TGG   1909
Ile Glu Ala Gly Glu Arg Phe Thr Val Tyr Val Val Val Pro Met Trp
                565                      570                      575
CCT GAG GGT GTT CCA GAG AGT GGA TCT GTT CAG GCA ATC CTG GAC TGG   1957
Pro Glu Gly Val Pro Glu Ser Gly Ser Val Gln Ala Ile Leu Asp Trp
            580                      585                      590
CAA AGG AGA ACA ATG GAG ATG ATG TAC ACT GAC ATT ACA GAG GCT CTC   2005
Gln Arg Arg Thr Met Glu Met Met Tyr Thr Asp Ile Thr Glu Ala Leu
        595                      600                      605
CAA GCC AAG GGA ATT GAA GCG AAC CCC AAG GAC TAC CTC ACT TTC TTC   2053
Gln Ala Lys Gly Ile Glu Ala Asn Pro Lys Asp Tyr Leu Thr Phe Phe
    610                      615                      620
TGC TTG GGT AAC CGT GAG GTG AAG CAG GCT GGG GAA TAT CAG CCT GAA   2101
Cys Leu Gly Asn Arg Glu Val Lys Gln Ala Gly Glu Tyr Gln Pro Glu
625                     630                      635                 640
GAA CAA CCA GAA GCT GAC ACT GAT TAC AGC CGA GCT CAG GAA GCT AGG   2149
Glu Gln Pro Glu Ala Asp Thr Asp Tyr Ser Arg Ala Gln Glu Ala Arg
                645                      650                      655
AGG TTC ATG ATC TAT GTC CAC ACC AAA ATG ATG ATA GTT GAC GAT GAG   2197
Arg Phe Met Ile Tyr Val His Thr Lys Met Met Ile Val Asp Asp Glu
```

```
             660                   665                   670
TAC ATC ATC ATC GGT TCT GCA AAC ATC AAC CAG AGG TCG ATG GAC GGC    2245
Tyr Ile Ile Ile Gly Ser Ala Asn Ile Asn Gln Arg Ser Met Asp Gly
         675                   680                   685
GCT AGG GAC TCT GAG ATC GCC ATG GGC GGG TAC CAG CCA TAC CAT CTG    2293
Ala Arg Asp Ser Glu Ile Ala Met Gly Gly Tyr Gln Pro Tyr His Leu
     690                   695                   700
GCG ACC AGG CAA CCA GCC CGT GGC CAG ATC CAT GGC TTC CGG ATG GCG    2341
Ala Thr Arg Gln Pro Ala Arg Gly Gln Ile His Gly Phe Arg Met Ala
705                 710                   715                   720
CTG TGG TAC GAG CAC CTG GGA ATG CTG GAT GAT GTG TTC CAG CGC CCC    2389
Leu Trp Tyr Glu His Leu Gly Met Leu Asp Asp Val Phe Gln Arg Pro
                725                   730                   735
GAG AGC CTG GAG TGT GTG CAG AAG GTG AAC AGG ATC GCG GAG AAG TAC    2437
Glu Ser Leu Glu Cys Val Gln Lys Val Asn Arg Ile Ala Glu Lys Tyr
            740                   745                   750
TGG GAC ATG TAC TCC AGC GAC GAC CTC CAG CAG GAC CTC CCT GGC CAC    2485
Trp Asp Met Tyr Ser Ser Asp Asp Leu Gln Gln Asp Leu Pro Gly His
        755                   760                   765
CTC CTC AGC TAC CCC ATT GGC GTC GCC AGC GAT GGT GTG GTG ACT GAG    2533
Leu Leu Ser Tyr Pro Ile Gly Val Ala Ser Asp Gly Val Val Thr Glu
    770                   775                   780
CTG CCC GGG ATG GAG TAC TTT CCT GAC ACA CGG GCC CGC GTC CTC GGC    2581
Leu Pro Gly Met Glu Tyr Phe Pro Asp Thr Arg Ala Arg Val Leu Gly
785                 790                   795                   800
GCC AAG TCG GAT TAC ATG CCC CCC ATC CTC ACC TCA TAGACGAGGA AGCACT 2633
Ala Lys Ser Asp Tyr Met Pro Pro Ile Leu Thr Ser
```

805 810

ACACTACAAT CTGCTGGCTT CTCCTGTCAG TCCTTCTGTA CTTCTTCAGT TTGGTGGCGA 2693

GATGGTATGG CCGTTGTTCA GAATTTCTTC AGAATAGCAG TTGTTACAGT TGTGAATCAT 2753

AAAGTAATAA GTGCAGTATC TGTGCATGGT TGAGTTGGGA AGAAGATCGG GGATGCAATG 2813

ATGCTTGTGA AGTTGTGATG CCGTTTGTAA GATGGGAAGT TGGGAACTAC TAAGTAATTG 2873

GCATGATTGT ACTTTGCACT ACTGTTTAGC GTTGTTGATA CTGGTTAACC GTGTGTTCAT 2933

CTGAACTTGA TTCTTGATGC AGTTTGTGGC ATTACCAGTT TATCATCGTT CTTCAGGAAA 2993

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAA 3040

SEQ ID NO.:2

SEQUENCE LENGTH: 812

SEQUENCE TYPE: AMINO ACID

MOLECULE TYPE: protein

ORIGINAL SOURCE

ORGANISM: Oryza sativa

SEQUENCE DESCRIPTION

Met Ala Gln Met Leu Leu His Gly Thr Leu His Ala Thr Ile Phe Glu
1 5 10 15

Ala Ala Ser Leu Ser Asn Pro His Arg Ala Ser Gly Ser Ala Pro Lys
20 25 30

Phe Ile Arg Lys Phe Val Glu Gly Ile Glu Asp Thr Val Gly Val Gly
35 40 45

Lys Gly Ala Thr Lys Val Tyr Ser Thr Ile Asp Leu Glu Lys Ala Arg
50 55 60

Val Gly Arg Thr Arg Met Ile Thr Asn Glu Pro Ile Asn Pro Arg Trp
65 70 75 80

Tyr Glu Ser Phe His Ile Tyr Cys Ala His Met Ala Ser Asn Val Ile

85 90 95

Phe Thr Val Lys Ile Asp Asn Pro Ile Gly Ala Thr Asn Ile Gly Arg

100 105 110

Ala Tyr Leu Pro Val Gln Glu Leu Leu Asn Gly Glu Glu Ile Asp Arg

115 120 125

Trp Leu Asp Ile Cys Asp Asn Asn Arg Glu Ser Val Gly Glu Ser Lys

130 135 140

Ile His Val Lys Leu Gln Tyr Phe Asp Val Ser Lys Asp Arg Asn Trp

145 150 155 160

Ala Arg Gly Val Arg Ser Thr Lys Tyr Pro Gly Val Pro Tyr Thr Phe

165 170 175

Phe Ser Gln Arg Gln Gly Cys Lys Val Thr Leu Tyr Gln Asp Ala His

180 185 190

Val Pro Asp Asn Phe Ile Pro Lys Ile Pro Leu Ala Asp Gly Lys Asn

195 200 205

Tyr Glu Pro His Arg Cys Trp Glu Asp Ile Phe Asp Ala Ile Ser Asn

210 215 220

Ala Gln His Leu Ile Tyr Ile Thr Gly Trp Ser Val Tyr Thr Glu Ile

225 230 235 240

Thr Leu Val Arg Asp Ser Asn Arg Pro Lys Pro Gly Gly Asp Val Thr

245 250 255

Leu Gly Glu Leu Leu Lys Lys Lys Ala Ser Glu Gly Val Arg Val Leu

260 265 270

Met Leu Val Trp Asp Asp Arg Thr Ser Val Gly Leu Leu Lys Arg Asp

275 280 285

Gly Leu Met Ala Thr His Asp Glu Glu Thr Glu Asn Tyr Phe His Gly

290 295 300

Ser Asp Val Asn Cys Val Leu Cys Pro Arg Asn Pro Asp Asp Ser Gly
305 310 315 320

Ser Ile Val Gln Asp Leu Ser Ile Ser Thr Met Phe Thr His His Gln
325 330 335

Lys Ile Val Val Val Asp His Glu Leu Pro Asn Gln Gly Ser Gln Gln
340 345 350

Arg Arg Ile Val Ser Phe Val Gly Gly Leu Asp Leu Cys Asp Gly Arg
355 360 365

Tyr Asp Thr Gln Tyr His Ser Leu Phe Arg Thr Leu Asp Ser Thr His
370 375 380

His Asp Asp Phe His Gln Pro Asn Phe Ala Thr Ala Ser Ile Lys Lys
385 390 395 400

Gly Gly Pro Arg Glu Pro Trp His Asp Ile His Ser Arg Leu Glu Gly
405 410 415

Pro Ile Ala Trp Asp Val Leu Tyr Asn Phe Glu Gln Arg Trp Arg Lys
420 425 430

Gln Gly Gly Lys Asp Leu Leu Leu Gln Leu Arg Asp Leu Ser Asp Thr
435 440 445

Ile Ile Pro Pro Ser Pro Val Met Phe Pro Glu Asp Arg Glu Thr Trp
450 455 460

Asn Val Gln Leu Phe Arg Ser Ile Asp Gly Gly Ala Ala Phe Gly Phe
465 470 475 480

Pro Asp Thr Pro Glu Glu Ala Ala Lys Ala Gly Leu Val Ser Gly Lys
485 490 495

Asp Gln Ile Ile Asp Arg Ser Ile Gln Asp Ala Tyr Ile His Ala Ile
500 505 510

Arg Arg Ala Lys Asn Phe Ile Tyr Ile Glu Asn Gln Tyr Phe Leu Gly

515 520 525

Ser Ser Tyr Ala Trp Lys Pro Glu Gly Ile Lys Pro Glu Asp Ile Gly

530 535 540

Ala Leu His Leu Ile Pro Lys Glu Leu Ala Leu Lys Val Val Ser Lys

545 550 555 560

Ile Glu Ala Gly Glu Arg Phe Thr Val Tyr Val Val Val Pro Met Trp

565 570 575

Pro Glu Gly Val Pro Glu Ser Gly Ser Val Gln Ala Ile Leu Asp Trp

580 585 590

Gln Arg Arg Thr Met Glu Met Met Tyr Thr Asp Ile Thr Glu Ala Leu

595 600 605

Gln Ala Lys Gly Ile Glu Ala Asn Pro Lys Asp Tyr Leu Thr Phe Phe

610 615 620

Cys Leu Gly Asn Arg Glu Val Lys Gln Ala Gly Glu Tyr Gln Pro Glu

625 630 635 640

Glu Gln Pro Glu Ala Asp Thr Asp Tyr Ser Arg Ala Gln Glu Ala Arg

645 650 655

Arg Phe Met Ile Tyr Val His Thr Lys Met Met Ile Val Asp Asp Glu

660 665 670

Tyr Ile Ile Ile Gly Ser Ala Asn Ile Asn Gln Arg Ser Met Asp Gly

675 680 685

Ala Arg Asp Ser Glu Ile Ala Met Gly Gly Tyr Gln Pro Tyr His Leu

690 695 700

Ala Thr Arg Gln Pro Ala Arg Gly Gln Ile His Gly Phe Arg Met Ala

705 710 715 720

Leu Trp Tyr Glu His Leu Gly Met Leu Asp Asp Val Phe Gln Arg Pro

725 730 735

Glu Ser Leu Glu Cys Val Gln Lys Val Asn Arg Ile Ala Glu Lys Tyr
740 745 750

Trp Asp Met Tyr Ser Ser Asp Asp Leu Gln Gln Asp Leu Pro Gly His
755 760 765

Leu Leu Ser Tyr Pro Ile Gly Val Ala Ser Asp Gly Val Val Thr Glu
770 775 780

Leu Pro Gly Met Glu Tyr Phe Pro Asp Thr Arg Ala Arg Val Leu Gly
785 790 795 800

Ala Lys Ser Asp Tyr Met Pro Pro Ile Leu Thr Ser
805 810

SEQ ID NO.:3

SEQUENCE LENGTH: 2804

SEQUENCE TYPE: NUCLEIC ACID

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE

ORGANISM: Zea maize

SEQUENCE DESCRIPTION

```
ATCACCGTCG TCATCAATCA CGGTGACCTT GCCCTCGCTG CCCGACTGGA ACCGGACGCT    60
GCTGCTGCTG GTAGGCTGAC AGCGAGGAGG ACGAGACGAG GGGGCC ATG GCT CAG      115
                                                   Met Ala Gln
                                                    1
ATC TTG CTC CAC GGC ACG CTC CAC GCC ACC ATC TTC GAG GCC GAG TCG     163
Ile Leu Leu His Gly Thr Leu His Ala Thr Ile Phe Glu Ala Glu Ser
      5                  10                  15
CTC TCC AAC CCG CAC CGC GCC ACT GGC GGC GCC CCC AAG TTC ATC CGC     211
Leu Ser Asn Pro His Arg Ala Thr Gly Gly Ala Pro Lys Phe Ile Arg
```

```
 20                  25                  30                  35
AAG CTT GTG GAA GGG ATC GAG GAC ACC GTG GGT GTC GGC AAG GGC GCC   259
Lys Leu Val Glu Gly Ile Glu Asp Thr Val Gly Val Gly Lys Gly Ala
                 40                  45                  50
ACC AAG ATA TAT GCC ACC GTC GAT CTC GAG AAG GCC CGT GTC GGG CGG   307
Thr Lys Ile Tyr Ala Thr Val Asp Leu Glu Lys Ala Arg Val Gly Arg
             55                  60                  65
ACC CGG ATG ATC TCC AAC GAG CCC GTG AAC CCT CGT TGG TAC GAG TCC   355
Thr Arg Met Ile Ser Asn Glu Pro Val Asn Pro Arg Trp Tyr Glu Ser
         70                  75                  80
TTC CAC ATC TAC TGC GCG CAC ATG GCC GCC GAC GTC ATC TTC ACC GTC   403
Phe His Ile Tyr Cys Ala His Met Ala Ala Asp Val Ile Phe Thr Val
     85                  90                  95
AAG ATC GAC AAC TCC ATC GGG GCC TCG CTC ATC GGG AGG GCC TAC TTG   451
Lys Ile Asp Asn Ser Ile Gly Ala Ser Leu Ile Gly Arg Ala Tyr Leu
100                 105                 110                 115
GCT GTC CAG GAC CTC CTG GGA GGG GAG GAG ATC GAC AAG TGG CTT GAA   499
Ala Val Gln Asp Leu Leu Gly Gly Glu Glu Ile Asp Lys Trp Leu Glu
                120                 125                 130
ATC TCC GAT GAA AAT CGT GAG CCT GTT GGG GAC AGC AAG ATC CAT GTG   547
Ile Ser Asp Glu Asn Arg Glu Pro Val Gly Asp Ser Lys Ile His Val
            135                 140                 145
AAG CTC CAG TAC TTT GAC GTC GGC AAG GAC CGT AAC TGG GCG AGG GGT   595
Lys Leu Gln Tyr Phe Asp Val Gly Lys Asp Arg Asn Trp Ala Arg Gly
        150                 155                 160
GTC CGG AGC ACC AAG TAC CCT GGT GTC CCT TAC ACC TTC TTC TCG CAG   643
Val Arg Ser Thr Lys Tyr Pro Gly Val Pro Tyr Thr Phe Phe Ser Gln
```

```
    165                 170                 175
AGG CAG GGG TGT AAG GTT ACT CTG TAC CAG GAC GCT CAT GTG CCG GAC    691
Arg Gln Gly Cys Lys Val Thr Leu Tyr Gln Asp Ala His Val Pro Asp
180                 185                 190                 195
AAC TTT GTT CCC AGG ATC CAG CTC GCT GAT GGC AAG AAC TAT GAG CCG    739
Asn Phe Val Pro Arg Ile Gln Leu Ala Asp Gly Lys Asn Tyr Glu Pro
                200                 205                 210
CAC AGG TGC TGG GAG GAT ATC TTT GAT GCT ATA AGC AAG GCT CAG CAT    787
His Arg Cys Trp Glu Asp Ile Phe Asp Ala Ile Ser Lys Ala Gln His
            215                 220                 225
TTG ATT TAC ATC ACT GGC TGG TCC GTG TAC ACA GAG ATC ACC TTG GTC    835
Leu Ile Tyr Ile Thr Gly Trp Ser Val Tyr Thr Glu Ile Thr Leu Val
        230                 235                 240
AGG GAC ACC AAC AGG CCA AAA CCT GGT GGT GAT GTT ACT CTT GGG GAG    883
Arg Asp Thr Asn Arg Pro Lys Pro Gly Gly Asp Val Thr Leu Gly Glu
    245                 250                 255
TTG CTC AAG AGG AAG GCC AGT GAA GGT GTC CGG GTG CTT ATG CTG GTG    931
Leu Leu Lys Arg Lys Ala Ser Glu Gly Val Arg Val Leu Met Leu Val
260                 265                 270                 275
TGG GAT GAC AGG ACT TCT GTC GGC CTG CTT AAG AAG GAT GGC TTG ATG    979
Trp Asp Asp Arg Thr Ser Val Gly Leu Leu Lys Lys Asp Gly Leu Met
                280                 285                 290
GCT ACC CAT GAT GAG GAG ACT GCA AAT TAC TTC CAT GGC ACG GAT GTC   1027
Ala Thr His Asp Glu Glu Thr Ala Asn Tyr Phe His Gly Thr Asp Val
            295                 300                 305
AAC TGT GTT CTG TGC CCT CGC AAC CCT GAT GAT TCT GGC AGC TTT GTC   1075
Asn Cys Val Leu Cys Pro Arg Asn Pro Asp Asp Ser Gly Ser Phe Val
```

```
         310                   315                   320
CAG GAT CTG CAG ATA TCA ACT ATG TTC ACG CAC CAC CAG AAG ATA GTA   1123
Gln Asp Leu Gln Ile Ser Thr Met Phe Thr His His Gln Lys Ile Val
     325                   330                   335
GTA GTC GAC CAT GAG ATG CCG AAC CAG GGA TCC CAG CAA AGG AGG ATA   1171
Val Val Asp His Glu Met Pro Asn Gln Gly Ser Gln Gln Arg Arg Ile
340                 345                   350                 355
GTC AGC TTC ATT GGT GGC ATT GAC CTT TGT GAT GGA AGA TAT GAT ACC   1219
Val Ser Phe Ile Gly Gly Ile Asp Leu Cys Asp Gly Arg Tyr Asp Thr
                360                   365                 370
CAG TAC CAC TCC TTG TTC AGG ACG CTT GAC ACT GTC CAT CAC GAT GAC   1267
Gln Tyr His Ser Leu Phe Arg Thr Leu Asp Thr Val His His Asp Asp
            375                   380                 385
TTC CAC CAG CCG AAC TTT GAG GGT GGG TCA ATC AAG AAA GGT GGC CCA   1315
Phe His Gln Pro Asn Phe Glu Gly Gly Ser Ile Lys Lys Gly Gly Pro
        390                   395                   400
AGG GAG CCA TGG CAT GAT ATC CAC TCA CGG TTG GAA GGG CCA ATC GCT   1363
Arg Glu Pro Trp His Asp Ile His Ser Arg Leu Glu Gly Pro Ile Ala
    405                   410                   415
TGG GAT GTT CTT TAC AAC TTT GAG CAG AGA TGG AGA AAG CAG GGT GGT   1411
Trp Asp Val Leu Tyr Asn Phe Glu Gln Arg Trp Arg Lys Gln Gly Gly
420                 425                   430                 435
42G GAC CTC CTT GTG CGT CTC AGG GAT CTT CCT GAC ATT ATC ATC CCC   1459
Lys Asp Leu Leu Val Arg Leu Arg Asp Leu Pro Asp Ile Ile Ile Pro
                440                   445                 450
CCT TCT CCT GTG ATG TTC CCG GAG GAC AGA GAG ACA TGG AAT GTT CAG   1507
Pro Ser Pro Val Met Phe Pro Glu Asp Arg Glu Thr Trp Asn Val Gln
```

```
                455                 460                 465
CTC TTC AGA TCC ATC GAT GGT GGT GCT GCT TTT GGC TTC CCC GAG ACT   1555
Leu Phe Arg Ser Ile Asp Gly Gly Ala Ala Phe Gly Phe Pro Glu Thr
            470                 475                 480
CCC GAG GAA GCT GCA AGA GCT GGG CTT GTG AGT GGA AAG GAT CAA ATC   1603
Pro Glu Glu Ala Ala Arg Ala Gly Leu Val Ser Gly Lys Asp Gln Ile
        485                 490                 495
ATC GAC CGG AGT ATC CAG GAT GCA TAC GTA AAC GCC ATA CGG AGG GCG   1651
Ile Asp Arg Ser Ile Gln Asp Ala Tyr Val Asn Ala Ile Arg Arg Ala
500                 505                 510                 515
AAG AAC TTC ATC TAC ATT GAG AAT CAG TAC TTC CTT GGA AGT TCA TAC   1699
Lys Asn Phe Ile Tyr Ile Glu Asn Gln Tyr Phe Leu Gly Ser Ser Tyr
                520                 525                 530
GGC TGG AAG CCC GAA GGC ATC AAG CCG GAA GAA ATC GGT GCT CTT CAC   1747
Gly Trp Lys Pro Glu Gly Ile Lys Pro Glu Glu Ile Gly Ala Leu His
            535                 540                 545
TTG ATT CCG AAG GAG CTC TCG CTG AAG ATT GTC AGC AAG ATT GAA GCT   1795
Leu Ile Pro Lys Glu Leu Ser Leu Lys Ile Val Ser Lys Ile Glu Ala
        550                 555                 560
GGG GAG CGG TTT ACT GTT TAT GTT GTG GTG CCA ATG TGG CCT GAG GGT   1843
Gly Glu Arg Phe Thr Val Tyr Val Val Val Pro Met Trp Pro Glu Gly
    565                 570                 575
GTT CCA GAA AGC GCT TCT GTT CAG GCA ATC CTT GAC TGG CAA AGG AGA   1891
Val Pro Glu Ser Ala Ser Val Gln Ala Ile Leu Asp Trp Gln Arg Arg
580                 585                 590                 595
ACG ATG GAG ATG ATG TAC ACT GAC ATC GCA CAA GCT CTC GAA GCC AAC   1939
Thr Met Glu Met Met Tyr Thr Asp Ile Ala Gln Ala Leu Glu Ala Asn
```

```
                    600                       605                       610
GGG ATT GAA GCA AAC CCC AAG GAC TAT CTC ACT TTC TTC TGC TTA GGT   1987
Gly Ile Glu Ala Asn Pro Lys Asp Tyr Leu Thr Phe Phe Cys Leu Gly
                615                       620                       625
AAC CGT GAG GTA AAG CAG GAG GGA GAA TAT GAA CCA GAG GAG CAC CCA   2035
Asn Arg Glu Val Lys Gln Glu Gly Glu Tyr Glu Pro Glu Glu His Pro
            630                       635                       640
GAA CCT GAC ACT GAT TAC ATC CGG GCT CAA GAG GCT AGG AGG TTT ATG   2083
Glu Pro Asp Thr Asp Tyr Ile Arg Ala Gln Glu Ala Arg Arg Phe Met
        645                       650                       655
ATC TAT GTT CAT ACC AAA ATG ATG ATA GTG GAC GAC GAG TAC ATC ATC   2131
Ile Tyr Val His Thr Lys Met Met Ile Val Asp Asp Glu Tyr Ile Ile
660                       665                       670                       675
ATT GGG TCC GCC AAC ATC AAC CAG AGG TCC ATG GAC GGT GCC AGG GAC   2179
Ile Gly Ser Ala Asn Ile Asn Gln Arg Ser Met Asp Gly Ala Arg Asp
                    680                       685                       690
TCC GAG ATC GCC ATG GGC GCG TAC CAG CCG TAC CAC TTG GCG ACT AGG   2227
Ser Glu Ile Ala Met Gly Ala Tyr Gln Pro Tyr His Leu Ala Thr Arg
                695                       700                       705
CAG CCT GCC CGG GGC CAG ATC CAT GGC TTC CGG ATG TCT CTT TGG TAC   2275
Gln Pro Ala Arg Gly Gln Ile His Gly Phe Arg Met Ser Leu Trp Tyr
            710                       715                       720
GAG CAC CTG GGA ATG CTG GAA GAC GTC TTC CAG CGG CCC GAG AGC GTA   2323
Glu His Leu Gly Met Leu Glu Asp Val Phe Gln Arg Pro Glu Ser Val
        725                       730                       735
GAG TGT GTG CAG AAG GTG AAC GAG GTC GCC GAG AAG TAC TGG GAC CTG   2371
Glu Cys Val Gln Lys Val Asn Glu Val Ala Glu Lys Tyr Trp Asp Leu
```

```
740                 745                 750                 755
TAC TCG AGC GAC GAC CTG GAG CAG GAC CTC CCG GGC CAC CTC CTC AGC   2419
Tyr Ser Ser Asp Asp Leu Glu Gln Asp Leu Pro Gly His Leu Leu Ser
                760                 765                 770
TAC CCG ATC GGT GTC ACT GCC GAC GGC AGC GTT ACC GAG CTG CCC GGG   2467
Tyr Pro Ile Gly Val Thr Ala Asp Gly Ser Val Thr Glu Leu Pro Gly
            775                 780                 785
ATG GAG AAC TTC CCC GAC ACC CGC GCC CGC GTC CTC GGG AAC AAG TCG   2515
Met Glu Asn Phe Pro Asp Thr Arg Ala Arg Val Leu Gly Asn Lys Ser
        790                 795                 800
GAT TAC CTC CCG CCC ATC CTC ACC ACA TAGAGTGCAC ACTGCAGGCA GCGCCAT 2569
Asp Tyr Leu Pro Pro Ile Leu Thr Thr
    805                 810
GGCTGCTCTC CTCTCTGGCC TCACCTTGGT GTCCCTGTGT TTGTGTTTGG GACACTGGAG 2629
GTTCAGATTG CAGTGTTGAT ATTATATCCC CCCTCCGTCC AGAGGGATTC GACGTTATTG 2689
AGTCATAATA AAATGCATTG TGCACGGTGG GAGACTGGGA GGATAGGAAT TATAGTTGTT 2749
TATTACAGTA CGACTGCTTA CTGCATCCAG ATTGTGTTGT CCCTAAAAAA AAAAA      2804
```

SEQ ID NO.:4

SEQUENCE LENGTH: 812

SEQUENCE TYPE: AMINO ACID

MOLECULE TYPE: protein

ORIGINAL SOURCE

ORGANISM: Zea maize

SEQUENCE DESCRIPTION

```
Met Ala Gln Ile Leu Leu His Gly Thr Leu His Ala Thr Ile Phe Glu
  1               5                  10                  15
```

```
Ala Glu Ser Leu Ser Asn Pro His Arg Ala Thr Gly Gly Ala Pro Lys
                20                  25                  30

Phe Ile Arg Lys Leu Val Glu Gly Ile Glu Asp Thr Val Gly Val Gly
            35                  40                  45

Lys Gly Ala Thr Lys Ile Tyr Ala Thr Val Asp Leu Glu Lys Ala Arg
        50                  55                  60

Val Gly Arg Thr Arg Met Ile Ser Asn Glu Pro Val Asn Pro Arg Trp
65                  70                  75                  80

Tyr Glu Ser Phe His Ile Tyr Cys Ala His Met Ala Ala Asp Val Ile
                85                  90                  95

Phe Thr Val Lys Ile Asp Asn Ser Ile Gly Ala Ser Leu Ile Gly Arg
            100                 105                 110

Ala Tyr Leu Ala Val Gln Asp Leu Leu Gly Gly Glu Glu Ile Asp Lys
        115                 120                 125

Trp Leu Glu Ile Ser Asp Glu Asn Arg Glu Pro Val Gly Asp Ser Lys
    130                 135                 140

Ile His Val Lys Leu Gln Tyr Phe Asp Val Gly Lys Asp Arg Asn Trp
145                 150                 155                 160

Ala Arg Gly Val Arg Ser Thr Lys Tyr Pro Gly Val Pro Tyr Thr Phe
                165                 170                 175

Phe Ser Gln Arg Gln Gly Cys Lys Val Thr Leu Tyr Gln Asp Ala His
            180                 185                 190

Val Pro Asp Asn Phe Val Pro Arg Ile Gln Leu Ala Asp Gly Lys Asn
        195                 200                 205

Tyr Glu Pro His Arg Cys Trp Glu Asp Ile Phe Asp Ala Ile Ser Lys
    210                 215                 220

Ala Gln His Leu Ile Tyr Ile Thr Gly Trp Ser Val Tyr Thr Glu Ile
```

225 230 235 240

Thr Leu Val Arg Asp Thr Asn Arg Pro Lys Pro Gly Gly Asp Val Thr

245 250 255

Leu Gly Glu Leu Leu Lys Arg Lys Ala Ser Glu Gly Val Arg Val Leu

260 265 270

Met Leu Val Trp Asp Asp Arg Thr Ser Val Gly Leu Leu Lys Lys Asp

275 280 285

Gly Leu Met Ala Thr His Asp Glu Glu Thr Ala Asn Tyr Phe His Gly

290 295 300

Thr Asp Val Asn Cys Val Leu Cys Pro Arg Asn Pro Asp Asp Ser Gly

305 310 315 320

Ser Phe Val Gln Asp Leu Gln Ile Ser Thr Met Phe Thr His His Gln

325 330 335

Lys Ile Val Val Val Asp His Glu Met Pro Asn Gln Gly Ser Gln Gln

340 345 350

Arg Arg Ile Val Ser Phe Ile Gly Gly Ile Asp Leu Cys Asp Gly Arg

355 360 365

Tyr Asp Thr Gln Tyr His Ser Leu Phe Arg Thr Leu Asp Thr Val His

370 375 380

His Asp Asp Phe His Gln Pro Asn Phe Glu Gly Gly Ser Ile Lys Lys

385 390 395 400

Gly Gly Pro Arg Glu Pro Trp His Asp Ile His Ser Arg Leu Glu Gly

405 410 415

Pro Ile Ala Trp Asp Val Leu Tyr Asn Phe Glu Gln Arg Trp Arg Lys

420 425 430

Gln Gly Gly Lys Asp Leu Leu Val Arg Leu Arg Asp Leu Pro Asp Ile

435 440 445

Ile Ile Pro Pro Ser Pro Val Met Phe Pro Glu Asp Arg Glu Thr Trp
450 455 460

Asn Val Gln Leu Phe Arg Ser Ile Asp Gly Gly Ala Ala Phe Gly Phe
465 470 475 480

Pro Glu Thr Pro Glu Glu Ala Ala Arg Ala Gly Leu Val Ser Gly Lys
485 490 495

Asp Gln Ile Ile Asp Arg Ser Ile Gln Asp Ala Tyr Val Asn Ala Ile
500 505 510

Arg Arg Ala Lys Asn Phe Ile Tyr Ile Glu Asn Gln Tyr Phe Leu Gly
515 520 525

Ser Ser Tyr Gly Trp Lys Pro Glu Gly Ile Lys Pro Glu Glu Ile Gly
530 535 540

Ala Leu His Leu Ile Pro Lys Glu Leu Ser Leu Lys Ile Val Ser Lys
545 550 555 560

Ile Glu Ala Gly Glu Arg Phe Thr Val Tyr Val Val Val Pro Met Trp
565 570 575

Pro Glu Gly Val Pro Glu Ser Ala Ser Val Gln Ala Ile Leu Asp Trp
580 585 590

Gln Arg Arg Thr Met Glu Met Met Tyr Thr Asp Ile Ala Gln Ala Leu
595 600 605

Glu Ala Asn Gly Ile Glu Ala Asn Pro Lys Asp Tyr Leu Thr Phe Phe
610 615 620

Cys Leu Gly Asn Arg Glu Val Lys Gln Glu Gly Glu Tyr Glu Pro Glu
625 630 635 640

Glu His Pro Glu Pro Asp Thr Asp Tyr Ile Arg Ala Gln Glu Ala Arg
645 650 655

Arg Phe Met Ile Tyr Val His Thr Lys Met Met Ile Val Asp Asp Glu

660 665 670

Tyr Ile Ile Ile Gly Ser Ala Asn Ile Asn Gln Arg Ser Met Asp Gly

675 680 685

Ala Arg Asp Ser Glu Ile Ala Met Gly Ala Tyr Gln Pro Tyr His Leu

690 695 700

Ala Thr Arg Gln Pro Ala Arg Gly Gln Ile His Gly Phe Arg Met Ser

705 710 715 720

Leu Trp Tyr Glu His Leu Gly Met Leu Glu Asp Val Phe Gln Arg Pro

725 730 735

Glu Ser Val Glu Cys Val Gln Lys Val Asn Glu Val Ala Glu Lys Tyr

740 745 750

Trp Asp Leu Tyr Ser Ser Asp Asp Leu Glu Gln Asp Leu Pro Gly His

755 760 765

Leu Leu Ser Tyr Pro Ile Gly Val Thr Ala Asp Gly Ser Val Thr Glu

770 775 780

Leu Pro Gly Met Glu Asn Phe Pro Asp Thr Arg Ala Arg Val Leu Gly

785 790 795 800

Asn Lys Ser Asp Tyr Leu Pro Pro Ile Leu Thr Thr

805 810

SEQ ID NO.:5

SEQUENCE LENGTH: 2799

SEQUENCE TYPE: NUCLEIC ACID

MOLECULE TYPE: genomic DNA

ORIGINAL SOURCE

ORGANISM: Zea maize

SEQUENCE DESCRIPTION

```
CAAGGGTGTA CATAGATTTG TCTCGTAAAA TAGTATTATA ATATTATAAA CTTATTACTC   60
TATCCGTTCT AAAATATAAG AACCTTATGA CTGGATGGAA CATTTCCTAG TACTACGAAT  120
CTGAACACAT GTCTAGATTC ATAGTACTAG GAAATGTCTC ATCGCGGTAC TAGGTTCTTA  180
TATTTTAGGA TGGAGGGAGT TTAATATAAA ACTAATGGTT AGAACTTTGA AAGTTTGATT  240
TTAAATGTCA AATATTTATG GCTGGAGGTA GTATAATATG TTTTTTTTGG GACGTAGACT  300
AGGTAGTATA ATATGTTTGG TTGTGTTTAG ATCCAATATT TGGATCCAAA CTTCAGTCAT  360
TTTCCATCAC ATCAACTTGT CATATACACA TAACTTTTCA GTCACATCAT CCCCAATTTC  420
AACCAAAATC AAACTTTGCG CTGAACTAAA CACAACCTTT GGGCCCGTTT AGTTCCCCAA  480
TTTTTTTCCC AAAAACATCA CATCGAATCT TTGGACACAT GCATGAAGCA TTAAATATAG  540
ATAAAAAGAA AAACTAATTG CACAGTTATG GAGGAAATCG CGAGACGAAT CTTTTAAGCC  600
TAATTAGTCC GTGATTAGCC ATAAGTGCTA CAGTAACCCA ATTGTGCTAA TGACGGCTTA  660
ATTAGTCTCC ACAAGATTCG TCTCGCAGTT TCCAGGCGAG TTCTGAAATT AGTTTTTTCA  720
TTCGTGTCCG AAAACCCCTT CCGACATCCG GTCAAACGTT CGATATGACA CCCACAAATT  780
TTCTTTTCCC CAACTAAACA CACCCTTTAT CTCTTACCCT CTGGCTCTTT CAGTAGGCAT  840
ATCCAAGACA GCTGGTAATG CAGGCTCGGA CATAATTTGA CAGTTACGTT CATGTGACCG  900
ACGGTTGATG CTAGTGCAAC TGCAACATAC TGTTCAGATG GATGTCCCAA CGAGCTCAAA  960
ACAACTTAGG TGGCGCGTCG CGATTCATCA ATAACTCAAA TGGAAGCGCA AGTGCACGTA 1020
CGAAAATGAC AGCGAGTGAG GTGGCGAGCC TCACCTTGGT GATCCCAACC GGATAAGCTA 1080
TGCATCAGCC AGTTTCGTGG GGCTGCACAT TTCGTCGAAC ACCTGGAGTC CACGCCGCCG 1140
GCGACGTCGG CACAGCGCGC CCGCCCACCG CCCACGCACG CGCTTGACTC CACCCATGTT 1200
CTCCCTTCTC GACGCCCGCG AAGCCAGCGA ACCGATCCGA GGAAGTCAAG CCCCCACCGC 1260
CACTTGGACC GACCTCGGGA CGACGACGCC CCCGCGCTCT TCTAGACGCG CGGACGACGC 1320
GGGCGCTGGC TCCGCGACGC GACGTCGCGG TCATGGAGTA ACCGCGACGG ACAGATACTT 1380
CTACCCGTTT TTAACCTCGC CTCCTCCTCC TCCCGGCTCG AGATCCGTGG CCACGACGCG 1440
TGGTGGGAAA CCGGGAACGA CGTGCACGCA CGCACACAGG GCAAGTTTCA GTAGAAAAAT 1500
CGCCGGCATC CAGATCGGGA CAGTCTCTCT TCTCCCGCAA TTTTATAATC TCGCTCGATC 1560
CAATCTGCTC CCCTTCTTCT TCTACTCTCC CCATCTCGGC TCTCGCCATC GCCATCCTCC 1620
```

TCTCCCTTCC CGGAGAAGAC GCCTCCCTCC GCCGATCACC ACCCGGTAAG CCCAGTGTGC 1680

TTAGGCTAAG CGCACTAGAG CTTCTTGCTC GCTTGCTTCT TCTCCGCTCA GATCTGCTTG 1740

CTTGCTTGCT TCGCTAGAAC CCTACTCTGT GCTGCGAGTG TCGCTGCTTC GTCTTCCTTC 1800

CTCAAGTTCG ATCTGATTGT GTGTGTGGGG GGGCGCAGGT AGGGCGAGGA GGGAGCCAAA 1860

TCCAAATCAG CAGCC ATG GCG CAG ATG CTG CTC CAT GGG ACG CTG CAC GCC 1911

Met Ala Gln Met Leu Leu His Gly Thr Leu His Ala

1 5 10

ACC ATC TTC GAG GCG GCG TCG CTC TCC AAC CCG CAC CGC GCC AGC GGA 1959

Thr Ile Phe Glu Ala Ala Ser Leu Ser Asn Pro His Arg Ala Ser Gly

15 20 25

AGC GCC CCC AAG TTC ATC CGC AAG GTTCGGACCC TTCTCCTTAA TCTACTCGTC 2013

Ser Ala Pro Lys Phe Ile Arg Lys

30 35

TTTGCTCTTG CTCTTTTTCT TTTGTGTCCC TTTCTTGTGT GTGCGTTTGC ATGAGCCCGA 2073

ATTTGATCTG CTAGTGCACA GTACAGTCAG ATACACTGAA ACGATCTGGA AATTCTGGAT 2133

TATTAGGAAA AATAAAGAGG TAGTAGACAA GAATTGGAGA TACTTTCTAT CAAGATTGGT 2193

CTATTATGCT TGGCCATTTC TTGTTTGACC CAAGTACTTC TTTGAATCTA GAGTTTGCTG 2253

TGTGTGATGT GGTGTGTGTT TGTGTCACCA AAAATCTTCA TTAGCTAAAA CTGAAATTTT 2313

ATTTATTAAC TGACCTACTA AAAATGTAGA GTTCTCTGTG TGTGATGTGT GCTTGTGTCA 2373

CCAAAAATCT TGATTTGATA GAGTTTTTAT TTATTTATTA ACTGACCTAC TACAAATCTA 2433

TTGCTGTATG CTATGTGTGT CTGTATCACC TGAAATGCAA TGTCTTCTTC TTTGTTGTTC 2493

TTGATCTAAC ACGTGAGCTC ATGTCAACAG TTT GTG GAG GGG ATT GAG GAC ACT 2547

Phe Val Glu Gly Ile Glu Asp Thr

40

GTG GGT GTC GGC AAA GGC GCC ACC AAG GTG TAT TCT ACC ATT GAT CTG 2595

Val Gly Val Gly Lys Gly Ala Thr Lys Val Tyr Ser Thr Ile Asp Leu

45 50 55 60

GAG AAA GCT CGT GTA GGG CGA ACT AGG ATG ATA ACC AAT GAG CCC ATC 2643

Glu Lys Ala Arg Val Gly Arg Thr Arg Met Ile Thr Asn Glu Pro Ile

65 70 75

AAC CCT CGC TGG TAT GAG TCG TTC CAC ATC TAT TGC GCT CAT ATG GCT 2691

Asn Pro Arg Trp Tyr Glu Ser Phe His Ile Tyr Cys Ala His Met Ala

80 85 90

TCC AAT GTG ATC TTC ACT GTC AAG ATT GAT AAC CCT ATT GGG GCA ACG 2739

Ser Asn Val Ile Phe Thr Val Lys Ile Asp Asn Pro Ile Gly Ala Thr

95 100 105

AAT ATT GGG AGG GCT TAC CTG CCT GTC CAA GAG CTT CTC AAT GGA GAG 2787

Asn Ile Gly Arg Ala Tyr Leu Pro Val Gln Glu Leu Leu Asn Gly Glu

110 115 120

GAG ATT GAC AGA 2799

Glu Ile Asp Arg

125

Claims

1. A cloned DNA which encodes phospholipase D originated from a plant.

2. The DNA according to claim 1, which encodes phospholipase D originated from a monocotyledonous plant.

3. A cloned DNA which encodes an amino acid sequence shown in SEQ ID NO. 2 or an amino acid sequence having the same sequence as shown in SEQ ID NO. 2 except that one or more amino acids are added, deleted or substituted, said amino acid sequence giving enzyme activity of phospholipase D.

4. The DNA according to claim 3, which encodes the amino acid sequence shown in SEQ ID NO. 2.

5. A cloned DNA which encodes an amino acid sequence shown in SEQ ID NO. 4 or an amino acid sequence having the same sequence as shown in SEQ ID NO. 4 except that one or more amino acids are added, deleted or substituted, said amino acid sequence giving enzyme activity of phospholipase D.

6. The DNA according to claim 5, whcih encodes the amino acid sequence shown in SEQ ID NO. 4.

7. The DNA according to claim 3, which has a nucleotide sequence shown in SEQ ID NO. 1 or has the same nucleotide sequence as shown in SEQ ID NO. 1 except that one or more nucleotides are added, deleted or substituted, said nucleotide sequence encodes an amino acid sequence giving enzyme activity of phospholipase D.

8. The DNA according to claim 7, which has a nucleotide sequence shown in SEQ ID NO. 1.

9. The DNA according to claim 5, which has a nucleotide sequence shown in SEQ ID NO. 3 or has the same nucleotide sequence as shown in SEQ ID NO. 3 except that one or more nucleotides are added, deleted or substituted, said nucleotide sequence encodes an amino acid sequence giving enzyme activity of phospholipase D.

10. The DNA according to claim 9, which has a nucleotide sequence shown in SEQ ID NO. 3.

11. The DNA according to claim 4, which has a nucleotide sequence from 182th to 2617th nucleotide in the nucleotide sequence shown in SEQ ID NO. 1.

12. The DNA according to claim 6, which has a nucleotide sequence from 107th to 2542th nucleotide in the nucleotide sequence shown in SEQ ID NO. 3.

13. A cloned DNA which regulates expression of phospholipase D gene originated from a plant.

14. A cloned DNA which has a nucleotide sequence shown in SEQ ID NO. 5 or has the same nucleotide sequence as shown in SEQ ID NO. 5 except that one or more nucleotides are added, deleted or substituted, said nucleotide sequence regulates expression of the DNA encoding an amino acid sequence giving enzyme activity of phospholipase D.

15. The DNA according to claim 14, which has a nucleotide sequence shown in SEQ ID NO. 5.

16. The DNA according to claim 5, which has a nucleotide sequence from 1st to 1875th nucleotide in the nucleotide sequence shown in SEQ ID NO. 5.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP94/01627

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl^6 C12N15/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl^5 C12N15/55

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CAS, BIOSIS, WPI, WPI/L

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	J. Biol. Chem. Vol. 269, No. 32 (1994), Xuemin Wang et al. "Cloning and Expression of Phosphatidylcholine-hydrolyzing Phospholipase D from Ricinus communis L", P. 20312-20317	1, 2, 13
X A	JP, A, 3-187382 (Toyo Jozo Co., Ltd.), August 15, 1991 (15. 08. 91) & EP, A, 435723	1, 2, 13 3-12, 14-16

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed
- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
December 6, 1994 (06. 12. 94)	December 27, 1994 (27. 12. 94)
Name and mailing address of the ISA/ Japanese Patent Office	Authorized officer
Facsimile No.	Telephone No.

Form PCT/ISA/210 (second sheet) (July 1992)